Synopsis

EXPLORING RIBOSOME MEDIATED TRANSLATION REGULATION IN MAMMALS

BY

NIVEDITA HARIHARAN

UNDER THE GUIDANCE OF

DR. DASARADHI PALAKODETI

INSTITUTE FOR STEM CELL SCIENCE AND REGENERATIVE MEDICINE

GKVK POST, BELLARY ROAD, BENGALURU - 560065

Synopsis

Exploring Ribosome mediated translation regulation in mammals

Introduction (*Chapter One of the thesis*)

Ribosomes, which are the hubs of protein synthesis (translation) have been typically viewed as passive components in translation regulation, with interest centred on intrinsic features of the mRNAs and *trans*-acting factors like RNA Binding Proteins (RBPs), non-coding RNAs. Basically composed of ribosomal RNAs (rRNAs) and proteins (RPs) forming two subunits (large subunit, LSU; small subunit, SSU), ribosomes possess domain-specific features. For instance, eukaryotic cytosolic ribosomes have nearly doubled in size than those seen in bacteria and archaea, which is mainly due to an increase in the number of RPs and the presence of multi-nucleotide insertions in eukaryotic rRNAs termed 'Expansion Segments' (ESs) (Veldman et al. 1981; Ware et al. 1983; Clark et al. 1984; Hassouna, Michot, and Bachellerie 1984; Brimacombe 1981). Therefore it is conceivable that changes in ribosomal protein composition and post-translational modifications, rRNA variations and modifications can give rise to a heterogenous pool of ribosomes, which can regulate and translate different subsets of mRNAs ('Ribosome Filter hypothesis') (Mauro and Edelman 2002). With the idea of 'ribosomal heterogeneity' gathering momentum of late, the traditional perception of ribosomes in a merely catalytic role in translation is changing. This notion is fairly nascent and still under debate, with only a handful of studies showing that ribosomes can indeed regulate translation (Díaz-López et al. 2019; Gómez Ramos et al. 2016; Fujii et al. 2018; Shankar et al. 2020).

Although first described at least four decades back, ESs are still a black box with very little known about their function. In the ribosome structure, ESs are present on the surface without perturbing the highly conserved core elements (Ben-Shem et al. 2011; Chandramouli et al. 2008a; Petrov et al. 2015; Anger et al. 2013a) and were initially considered to be non-functional insertions. These multi-nucleotide insertions of variable length are present at conserved positions along the eukaryotic rRNAs when compared to the bacterial rRNAs. While ESs of the SSU are of similar lengths across eukaryotes, those of the LSU can be variable in length. For instance, a LSU ES called ES7L is only around 100 nucleotides in *Saccharomyces cerevisiae* while it is more than 800 nucleotides in *Homo sapiens* (Parker et al. 2015). But not all LSU ESs necessarily show a size increase that correlates with organismal complexity. Further, even though 14 ESs from the SSU and 41 ESs from the LSU have been reported overall, some of them could be specific to subsets of clades (Yokoyama and Suzuki 2008). The species-specific diversification of ESs and their role in regulating translation has largely remained unaddressed. The nature and known potential functions of ESs are discussed in depth in our review "The story of rRNA expansion segments: Finding functionality amidst diversity" published in WIREs RNA with DOI: https://doi.org/10.1002/wrna.1732 To understand the variability of the different LSU ESs

across eukaryotes, we compared 28S rRNA sequences from species representing various clades across eukaryotes. Given that the size increase of the SSU rRNA is not as drastic among eukaryotes as that of the LSU rRNA, we decided to focus on the latter. From our comparison, we observed a LSU ES called ES30L, specific to endothermic vertebrates (mammals and birds). ES30L is highly GC-rich in *H. sapiens*, very flexible, and positioned near the ribosome's t-RNA and mRNA exit site. These aspects prompted us to focus on this rRNA segment and study its interactome, which could help us gain insights into its function.

Therefore, the primary objectives of our study have been -

1. To compare ES30L across eukaryotes.
2. To examine the potential of ES30L to bind to protein-coding transcripts in *H. sapiens*.
3. To investigate the protein interactome of ES30L in *H. sapiens*.

Materials and Methods (*Chapter Two of the thesis*)

The materials and methods used in this study are elaborated in the manuscript "ES30L: A large subunit rRNA expansion segment enriched in mammals and birds" published in bioRxiv with the DOI: https://doi.org/10.1101/2022.09.30.510333. This manuscript is currently under peer review.

Results

The results of this study will be elaborated over five chapters as mentioned below.

Comparison of ESs across eukaryotes (*Chapter Three of the thesis*)

We mined 28S rRNA sequences from public databases like Silva LSU r132 (26) and NCBI Nucleotide and constructed a multiple sequence alignment (MSA) using them. The major ESs in the human LSU rRNAs were marked in the MSA based on the coordinates taken from Parker et al., (Table 1) and were used to extract the concurrent regions from all the other species used in the MSA. We also mapped these expansions on the structure of the human ribosome. Our observations from this analysis are summarized below -

1. ESs are present on the ribosomal surface in both subunits (Figure 1)
2. ESs of the LSU such as ES5L, ES24L and ES27L show variability in length and/or base composition across bilaterians.
3. Other ESs such as ES9L, ES12L, ES26L are similar in length and base composition.
4. Some expansions contain insertions that are specific to a clade(s). For instance, ES30L, ES39L and ES15L have insertions present only in endothermic vertebrates. Arthropods have some expansions that are unique to their clade such as those in ES31L. Some expansions are slightly longer in invertebrates than in vertebrates such as ES19L and ES20L.

The function of almost all of these expansions and the relevance of species-specific extensions are unknown. We focussed on ES30L, a highly GC-rich ES with mammal and bird-specific expansion that is an extension of L1-stalk.

Table 1 : Co-ordinates of the LSU ESs in *H. sapiens*.

ES	Co-ordinates			GC%
	U13369_7935-12969 (taken from Parker et al., 2015)	**Mapping in PDB:4UG0**	**Mapping in NR_003287.4**	
ES5L	114-156	114-159	114-159	78.57
ES7L	465-1265	474-1278	474-1278	83.75
ES9L	1384-1487	1397-1500	1397-1500	80.58
ES10L	1682-1711	1695-1724	1695-1724	75.86
ES12L	1793-1829	1806-1842	1806-1842	77.78
ES15L	2075-2256	2088-2269	2088-2269	83.98
ES19L	2439-2494	2452-2507	2452-2507	74.55
ES20L	2538-2575	2551-2588	2551-2588	70.27
ES24L	2685-2705	2698-2718	2698-2718	85
ES26L	2751-2764	2764-2777	2764-2777	61.54
ES27L	2875-3586	2888-3607	2888-3607	86.64
ES30L	3955-4013	3976-4035	3976-4035	91.8
ES31L	4057-4124	4079-4154	4079-4154	83.58
ES39L	4698-4905	4729-4940	4729-4940	82.61
ES41L	4983-4996	5018-5031	5018-5031	69.23

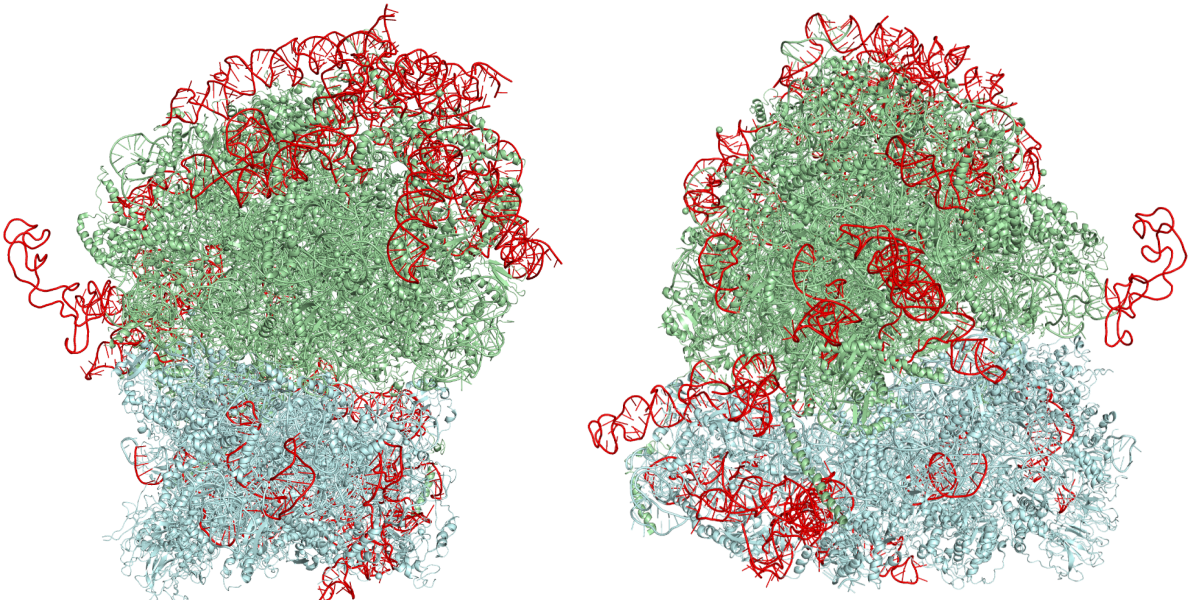

Figure 1: This figure shows the human ribosome structure (PDB 4UG0) with all the major ESs marked in red, LSU in green and SSU in cyan. ***(A)*** Front View ***(B)*** Back View

ES30L is a large subunit ES specific to endothermic vertebrates (*Chapter Four of the thesis*)

To compare the ES30L region across eukaryotes, we extracted the region corresponding to nucleotides 3975-4035 in the human 28S rRNA sequence from the MSA (Figure 2A). We also manually edited this section of the alignment to make it better suited to analyse the extent of conservation. This region showed the highest expansion in mammals (Figure 2C) and was highly GC-rich (~92% in *H. sapiens*). We observed that amphibians and arthropods also had short expansions in this region, but exhibited high variability in length within these two groups. Further, we also observed that while ES30L was GC-rich in chordates, it had among the lowest GC-richness in arthropods (Figure 2D). We used the edited alignment to calculate Gap-Adjusted Shannon's Entropy (GASE), which is plotted in Figure 2E and noted that even though ES30L was well conserved within both clades, the conservation was better among mammals than birds. Mapping this region on the structure of the human ribosome showed that ES30L extrudes as a disordered stretch from the L1-stalk and is usually not completely resolved in cryo-EM because of its tentacle-like nature (Anger et al. 2013a; Chandramouli et al. 2008a; Khatter et al. 2015; Natchiar et al. 2017). It is usually reconstructed with the help of secondary structure modelling (Figure 2B, 1C). Given its GC-richness and position on the ribosome, we hypothesized that it could bind to mRNAs and proteins and regulate translation.

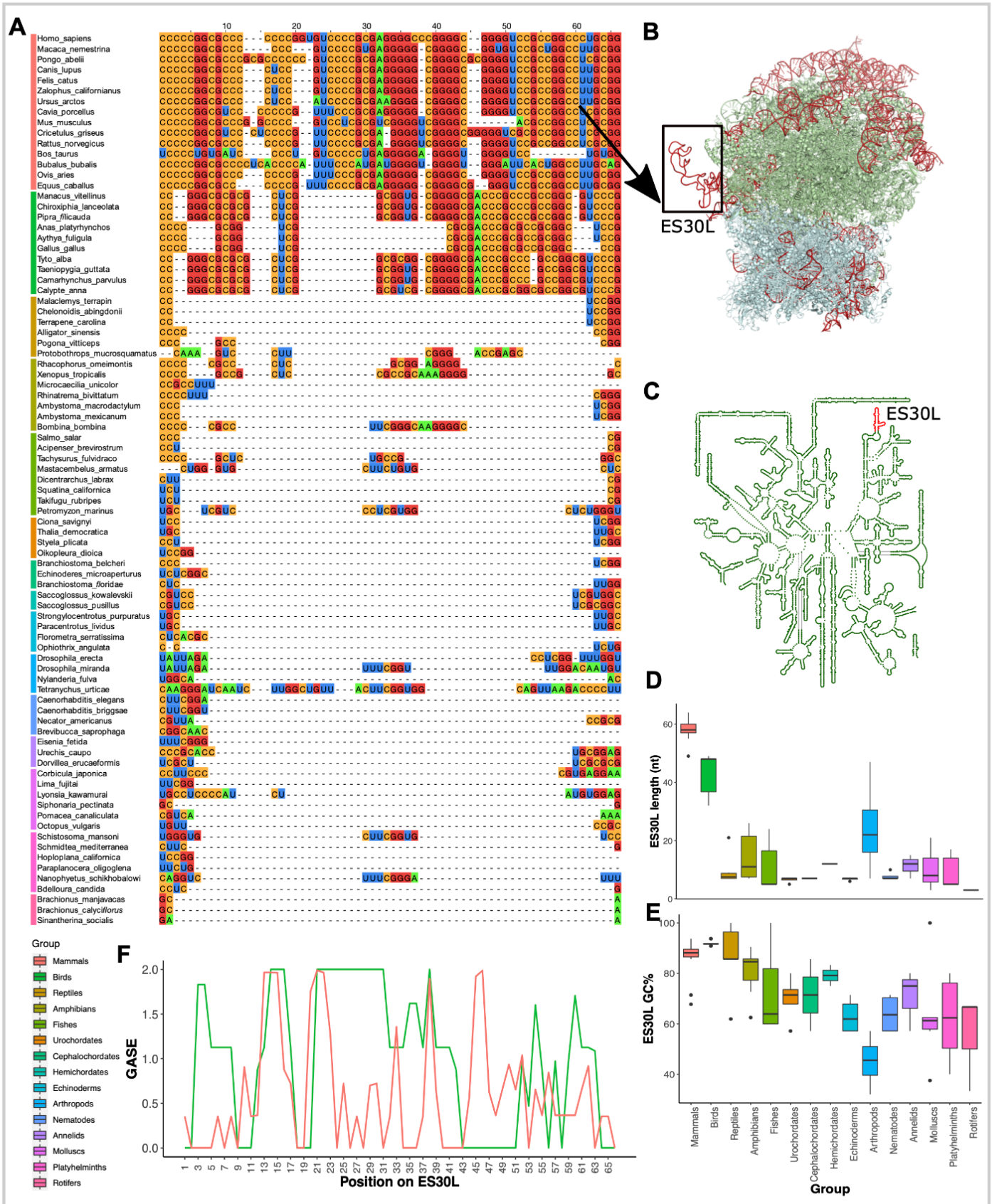
A
B
ES30L
C
ES30L
D
ES30L length (nt)
E
ES30L GC%
Group
Mammals
Birds
Reptiles
Amphibians
Fishes
Urochordates
Cephalochordates
Hemichordates
Echinoderms
Arthropods
Nematodes
Annelids
Molluscs
Platyhelminths
Rotifers
F
GASE
Position on ES30L

Figure 2: ES30L is a 28S rRNA expansion segment that is expanded largely in mammals and birds. This figure shows the comparison of the ES30L region from different clades (denoted by the coloured bars on the left of the MSA) across bilateria

(A) Multiple Sequence Alignment (MSA) of the ES30L region from 85 organisms across Bilateria.

(B) Front view of the human 80S ribosome (PDB: 4UG0 10.2210/pdb4UG0/pdb) (Khatter et al. 2015). The boxed region on the ribosome is ES30L and corresponds to the rRNA stretch shown in the MSA. It is important to note that the tertiary structure of ES30L is modelled based on both secondary structure prediction and cryo-EM data and is not completely resolved because of its flexibility. The ribosome structure was rendered using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

(C) Secondary structure of the *H. sapiens* 28S rRNA with the ES30L highlighted in red, created using the RiboVision web server (Bernier et al. 2014). The coordinates of ES30L and the other ESs were taken from Parker *et al*., (Parker et al. 2015).

(D) & (E) Box plots displaying the lengths and the GC% of the ES30L region among the sampled organisms across bilateria.

(F) Gap Adjusted Shannon's Entropy (GASE) plot for the ES30L region from mammalian (red) and bird (green) sequences shown in the MSA.

ES30L possesses complementarity to many protein-coding transcripts (*Chapter Five of the thesis*)

We developed a computational pipeline to probe the extent of complementarity prevalent between ES30L and protein-coding transcripts from *H. sapiens*. For this analysis, we used a 'sliding window' approach and extracted continuous overlapping 7 to 15 nucleotide stretches from ES30L. Transcript regions that exhibited complementarity (with only canonical Watson-Crick pairing) to these ES30L stretches either with zero or one non-terminal mismatch were considered for further analysis. Regions other than ESs in the 28S rRNA (termed as core segments or CSL) were used as a control to delineate the specificity of ES30L to transcripts. Our key observations from this set of analyses are :

1. Overall, ES30L had 7.25% higher pooled complementary hits with no mismatch to transcripts than CSL (paired Wilcoxon signed-rank test, $p < 0.05$; Figure 3A), with 84.54% of the transcripts possessing at least one stretch perfectly complementary to ES30L.
2. ES30L exhibited 36.12% more complementarity than CSL at lengths of 9 nucleotides or higher (Figure 3B). This was even higher when one mismatch was allowed, as more than 50% of the stretches with lengths of 7 or 8 got extended to 9 nucleotides or longer.
3. Complementary transcript stretches that are longer than 15 nucleotides were sparse with 1148 hits from 560 genes if one mismatch was allowed. This gene set included many that are involved in developmental pathways and neurogenesis among others.
4. Around two-thirds of the complementarity observed between ES30L and transcripts stemmed from the latter half of the ES (after position 25; Figure 3C), with the skew more

pronounced for longer complementary stretches. With CSL, the complementarity observed was uniform across the segments (Figure 3D).

5. The density of complementary stretches in the transcripts was nearly three times higher in 5' UTR when compared to CDS and 3' UTR (Figure 3A), with an enrichment around the start codon (Figure 3E-F). The differential density among the transcript sectors was not stark in the case of CSL.
6. Many transcripts possessed more than one stretch that is complementary to ES30L, with 425 of them containing more than 50 such stretches. Gene ontology analysis of this set showed an enrichment of genes involved in neuronal development and differentiation.

All the trends observed here hold true even when one mismatch was allowed in the complementarity analysis. Our results so far suggest that ES30L has a higher propensity to interact with transcripts, especially around the start codon.

A

Category	No. Of Complementary Hits	
	ES30L	CSL
All mRNA sectors	525190	489677
5' UTR (per 100nt)	1.0347	0.4759
CDS (per 100nt)	0.3437	0.2789
3' UTR (per 100nt)	0.2128	0.2826

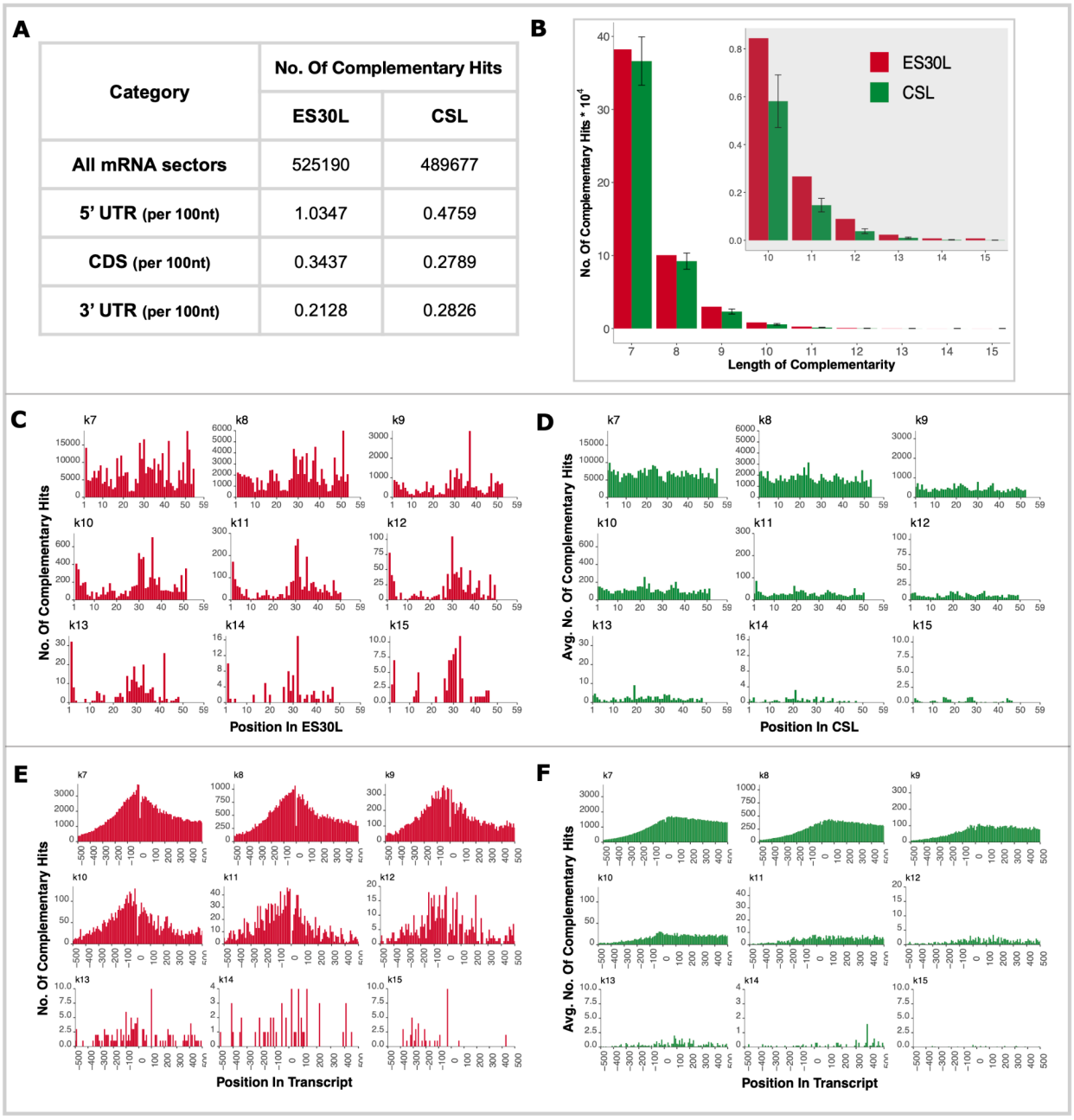

Figure 3: ES30L (from human 28S rRNA) possesses complementarity to many protein-coding human transcripts. This figure shows the summarized comparison between ES30L and CSL (core segment from the LSU; see *Methods*) regions with regard to their contiguous complementarity (zero mismatch) to protein-coding transcripts. In the case of CSL segments, all the data shown here is the average value from the ten segments used in the analysis.

(A) Table showing the total number of complementary hits in all mRNA sectors (5' UTR, CDS, 3' UTR) when matched to ES30L and CSL segments. The table also shows the density of complementary hits from each mRNA sector, which is the total complementary hits from that sector divided by the cumulative length of the sector from all the transcripts.

(B) Bar plot showing the number of complementary stretches ranging from a length of 7 to 15 or more nucleotides for ES30L (red) and CSL (green) fragments. The inset in this panel is the data for 10 nucleotides and higher, magnified for better clarity.
(C) & ***(D)*** Histogram showing the number of complementarity stretches to transcripts, arising from each position on ES30L (red) and CSL (green) fragments respectively.
(E) & ***(F)*** Histogram showing the number of complementarity stretches arising from each bin (bin size used here is 10) on the protein-coding transcripts, with ES30L (red) and CSL (green) fragments. The '0' on the x-axis denotes the start codon (TSS) and the negative numbers upstream of it, denoting positions in the 5' UTR. Only the 500 bases flanking the TSS have been shown in these plots. The complete histogram of this data is provided in Supplementary Figures S2-IC, S2-ID, S2-IIE, and S2-IIF.

ES30L could bind to protein-coding transcripts (*Chapter Six of the thesis*)

To test whether ES30L can interact with transcripts, we pulled down mRNAs from HEK293T cell lysates using a biotinylated ES30L mimic followed by high-throughput RNA sequencing. We used a random oligonucleotide of the same length as ES30L as a control in the *in vitro* pulldown assay. Concurrently, we re-analysed a published RNA-RNA interactome dataset (Aw et al. 2016) from HeLa cells to probe for ES30L binding to transcripts *in vivo*. Data from these analyses was overlaid with our bioinformatic predictions to check the extent of complementarity in these transcripts. Our results are summarized below :

1. Our sequencing data post the *in vitro* pulldown showed that 1550 transcripts were more than two-fold enriched ($q<0.05$) in the ES30L fraction over input (Figure 4A). Out of these, 100 transcripts showed a 50% higher fold enrichment in the ES30L fraction than CSL, when compared to input (Figure 4B). Gene set enrichment analysis revealed that many of these transcripts are involved in various aspects of RNA metabolism and translation (Figure 4C).
2. More than 90% of the transcripts selectively enriched in the ES30L fraction from our pulldown included at least one stretch (of any length) complementary to ES30L either with or without a mismatch.
3. From the *in vivo* RNA-RNA interactome dataset, we retrieved 193 inter and intra-molecular interactions involving the ES30L region, out of which 143 were with protein-coding transcripts (Figure 4D). Nearly 65% of the 143 interacting transcript stretches encompass at least one complementary stretch to ES30L (Figure 4E).
4. The 143 *in vivo* interactions involve 118 genes, 25 of which are also present in our *in vitro* pulldown data. This overlap could vary because these data are from two different cell lines.
5. Since interactions involving RNAs can also be structure based, we further probed for putative IRES elements (Yang et al. 2021; Zhao et al. 2020) in the ES30L interacting mRNA stretches from the *in vivo* interactome data. Genes such as RPS11, FTH1 and

NHP2 include ES30L-binding stretches which are part of predicted IRES elements. These regions have multiple short stretches with complementarity to ES30L.

6. Further, we did *in silico* analysis for putative structural motifs in 5' UTR stretches containing complementarity to ES30L from the transcripts in our pulldown. The results indicated that around 4-14% of the transcripts could form consensus motifs (Figure 4F), which contain binding motifs for proteins such as HNRNPs, SRSFs, DHX9, PTBP. Interestingly, binding motifs for these proteins are also present in ES30L.

Based on our observations so far, we hypothesize that ES30L could bind to complementary transcript regions, some of which may be part of IRES structures.

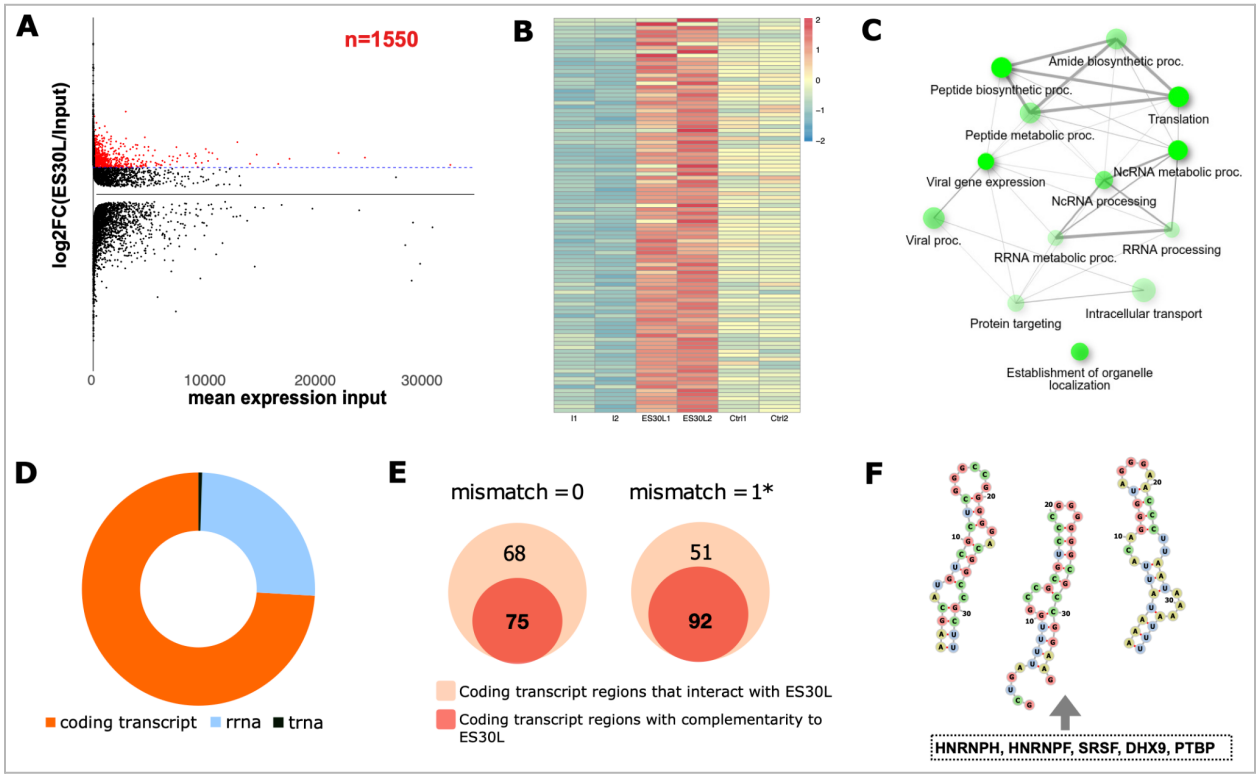

Figure 4: ES30L can potentially interact with protein-coding human transcripts. This figure summarises data that suggests that such an interaction is possible.

(A) MA plot showing the distribution of the fold change of all the transcripts between the ES30L fraction and the input, relative to the mean expression in the input. The points highlighted in red represent the 1550 transcripts that are more than two-fold enriched ($q<0.05$) in the ES30L fraction over the input.

(B) Heatmap showing the expression profiles (normalised read count) of the 100 transcripts that are more than two-fold enriched in the ES30L fraction over the input fraction and 50% higher fold change than the random fragment in the pulldown.

(C) Network plot showing the gene categories over-represented in the 100 transcripts that are more than two-fold enriched in the ES30L fraction over the input fraction and 50% higher fold change than the random fragment in the pulldown.

(D) Pie chart showing the distribution of the RNAs that interact with the ES30L region reported in Aw *et. al. (Aw et al. 2016)*.
(E) Venn diagram showing the overlap between the number of transcript regions interacting with the ES30L region reported in Aw *et. al. (Aw et al. 2016)* and the number of those transcript regions harbouring at least one stretch complementary to ES30L.
(F) RNA motifs obtained from *de novo* secondary structure prediction. The sequences used in this prediction were taken from the 100 transcripts from our pull-down and were 5' UTR regions which include stretches complementary to ES30L flanked by 20 nucleotides.

ES30L can interact with various RNA Binding Proteins (*Chapter Seven of the thesis*)

To identify the proteins that interact with ES30L, we did an *in vitro* pulldown of proteins from HEK293T cells using the same biotinylated RNA mimics as before, followed by mass spectrometry. We also re-analysed published CLIP-seq datasets (Van Nostrand et al. 2020; Porter et al. 2021) from K562 or HCT116 cells for proteins that were selectively enriched in the ES30L fraction in our pulldown to probe their interaction with ES30L *in vivo*. The results are summarized below -

1. In total fifty-eight proteins (Figure 5A) were more than 1.5 fold enriched ($q<0.05$) in the ES30L fraction when compared to input, which included various categories of RBPs such as a)Ribosome biogenesis factors (NKRF, NCL) b)Helicases (DDX21, DHX30, DHX9) c)G-quadruplex binding proteins (hnRNPF) d)Splicing factors (SRSF1, SRSF2, SRSF6, U2AF2) e)Ribosomal proteins f)Proteins involved in translation (PCBP1, HNRNPK, IGF2BP1, SRP14, RBM4).
2. Overall, gene ontology analysis of these proteins showed that many of them are involved in various aspects of RNA metabolism (Figure 5B).
3. Out of the 58 identified proteins, 42 were part of a published ribo-interactome (Simsek et al. 2017)(Figure 6A).
4. In our analysis of the CLIP-seq data DHX30, NPM1 showed enrichment in reads mapping to ES30L relative to the input, suggesting that they could potentially interact with ES30L *in vivo*. Other proteins like PCBP1, SRSF1, IGF2BP1, U2AF2 or HNRNPK did not show such enrichment in this dataset (Figure 6B).
5. Since *in vivo* molecular interactions could be influenced by their micro-environment, we compared CLIP-Seq data for PCBP1 from two cell types (K562 and HCT116). Our analysis indicated a preferential mapping of reads towards the latter half of ES30L in one cell type and not the other, hinting at the possibility that interactions of RBPs with ES30L could be context-dependent (Figure 6C).

Based on our *in silico* analysis, many proteins from our *in vitro* pulldown contain multiple RNA binding domains and have motifs on both ES30L and complementary transcript stretches. Therefore, interactions between ES30L and transcripts may be mediated by RNA-binding proteins.

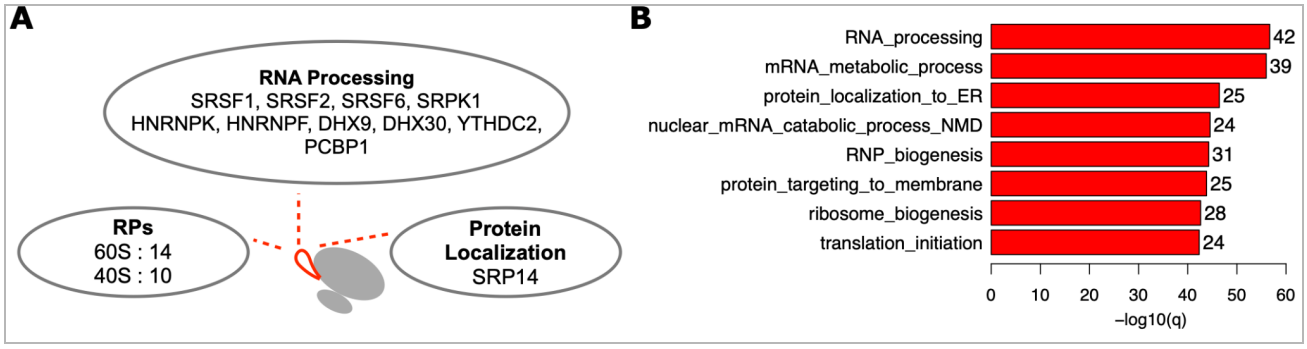

Figure 5: This figure gives an overview of the potential protein interactome of ES30L
(A) Illustration depicting the broad protein categories that were enriched more than two-fold (q<0.05) over the total protein fraction in our pulldown.
(B) Bar plot showing the GO (biological process) categories of the proteins that were enriched more than two-fold (q<0.05) over the total protein fraction in our pulldown.

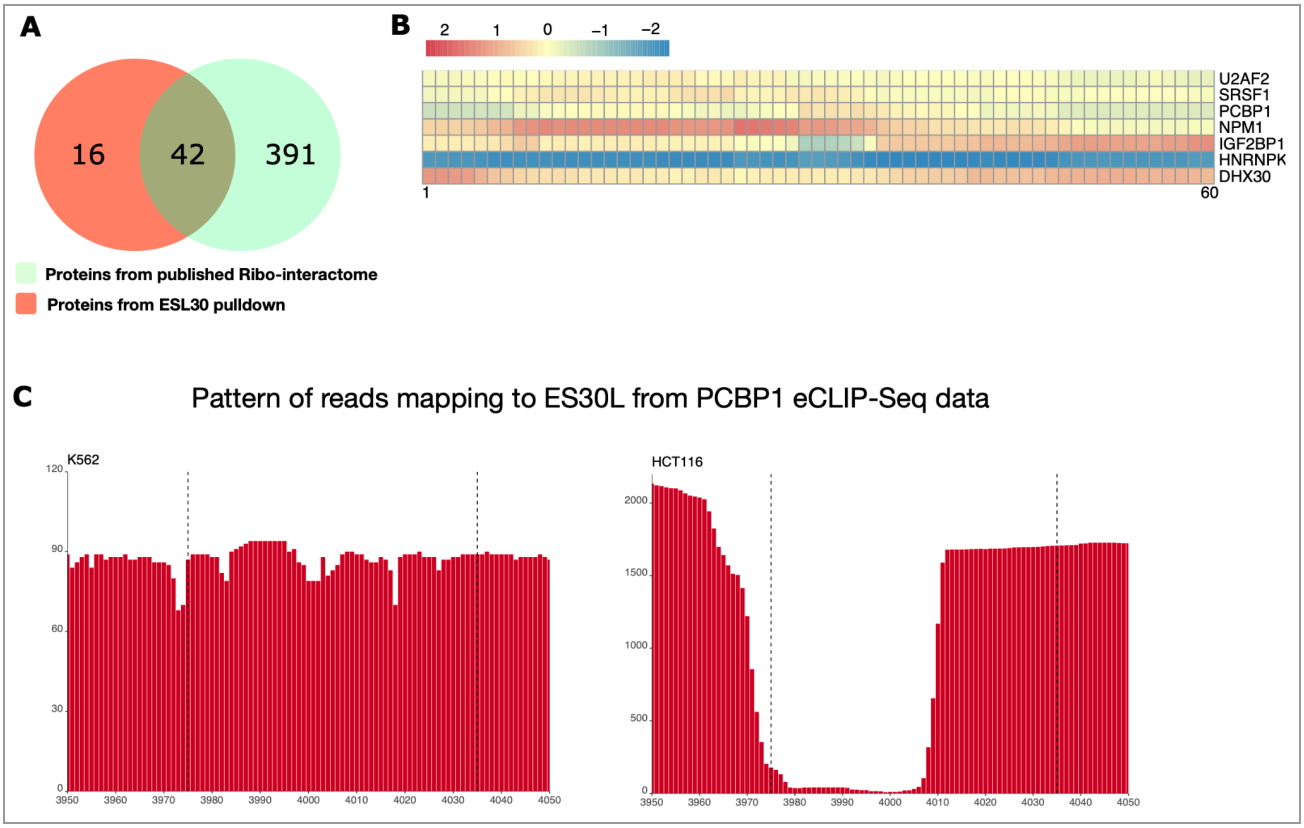

Figure 6: ES30L could interact with RBPs.
(A) Venn diagram showing the overlap between the proteins identified in our pulldown experiment and a published ribo-interactome ((Shi et al. 2017)).
(B) Heatmap showing the enrichment of reads (for the protein fraction over their size-matched input) in the ES30L region from published CLIP-seq data ((Van Nostrand et al. 2020)) for eight proteins that were identified in our pulldown.

(C) & (D) Coverage plot showing the read depth pattern across the ES30L stretch (U13369:3955-4013; Supplementary Table S1C) from two different studies - Van Nostrand et al. 2020 (*left*; K562 cell line) and Porter et al. 2021 (*right*; HCT116 cell line).

Conclusion and Future Directions (*Chapter Eight of the thesis*)

Some of the key findings from my work are:

1. ESs (especially of the LSU) may not be uniform (in terms of length and base composition) across eukaryotes and may contain species-specific variations.
2. ES30L has expanded the most in mammals, with the second-highest expansion in birds. It is well conserved in both these clades.
3. This segment has a high degree of complementarity to several protein-coding mRNAs, which are highly clustered around the start codon, with the highest density in the 5' UTR.
4. Our *in vitro* pulldown combined with our analysis of other data shows that ES30L can interact with mRNAs, which could involve base-pairing and structural 5' UTR motifs.
5. A similar *in vitro* pulldown of proteins indicates that ES30L could bind proteins that are involved in various aspects of RNA metabolism. Interestingly, some of these proteins have been shown to be important for cap-independent translation and based on our *in silico* analysis, also possess binding motifs on both ES30L and 5' UTR stretches in mRNAs from our RNA pulldown.

In summary, we think that ES30L can interact with secondary structural elements in the 5' UTR, RBPs and influence translation initiation, although such an influence could be either stabilising or inhibitory. This is one possible role of ES30L, though it could be involved in other cellular processes too. Our hypothesis is very difficult to test *in cellulo* using a knock-out strategy because rRNA is present as tandem repeats across multiple chromosomes in the human genome. Future strategies to establish the function of ES30L *in vivo* could be to design highly specific ES30L inhibitors that prevent all its interactions or knock-in the segment in to rRNA from other species. Considering that nothing is known about most ESs, we think it is important to develop techniques that are tailored to these enigmatic segments, as it could widen our understanding of ribosomes in general.

References

1. Anger, Andreas M., Jean-Paul Armache, Otto Berninghausen, Michael Habeck, Marion Subklewe, Daniel N. Wilson, and Roland Beckmann. 2013a. "Structures of the Human and Drosophila 80S Ribosome." *Nature* 497 (7447): 80–85.
2. Aw, Jong Ghut Ashley, Yang Shen, Andreas Wilm, Miao Sun, Xin Ni Lim, Kum-Loong Boon, Sidika Tapsin, et al. 2016. "In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation." *Molecular Cell* 62 (4): 603–17.
3. Ben-Shem, Adam, Nicolas Garreau de Loubresse, Sergey Melnikov, Lasse Jenner, Gulnara

Yusupova, and Marat Yusupov. 2011. “The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution.” *Science* 334 (6062): 1524–29.

4. Bernier, Chad R., Anton S. Petrov, Chris C. Waterbury, James Jett, Fengbo Li, Larry E. Freil, Xiao Xiong, et al. 2014. “RiboVision Suite for Visualization and Analysis of Ribosomes.” *Faraday Discussions* 169 (May): 195–207.
5. Brimacombe, Richard. 1981. “Secondary Structure and Evolution of Ribosomal RNA.” *Nature*. https://doi.org/10.1038/294209a0.
6. Chandramouli, Preethi, Maya Topf, Jean-François Ménétret, Narayanan Eswar, Jamie J. Cannone, Robin R. Gutell, Andrej Sali, and Christopher W. Akey. 2008a. “Structure of the Mammalian 80S Ribosome at 8.7 A Resolution.” *Structure* 16 (4): 535–48.
7. Clark, C. G., B. W. Tague, V. C. Ware, and S. A. Gerbi. 1984. “Xenopus Laevis 28S Ribosomal RNA: A Secondary Structure Model and Its Evolutionary and Functional Implications.” *Nucleic Acids Research* 12 (15): 6197–6220.
8. Díaz-López, Irene, René Toribio, Juan José Berlanga, and Iván Ventoso. 2019. “An mRNA-Binding Channel in the ES6S Region of the Translation 48S-PIC Promotes RNA Unwinding and Scanning.” *eLife* 8 (December). https://doi.org/10.7554/eLife.48246.
9. Fujii, Kotaro, Teodorus Theo Susanto, Saumya Saurabh, and Maria Barna. 2018. “Decoding the Function of Expansion Segments in Ribosomes.” *Molecular Cell* 72 (6): 1013–20.e6.
10. Gómez Ramos, Lizzette M., Johanna M. Smeekens, Nicholas A. Kovacs, Jessica C. Bowman, Roger M. Wartell, Ronghu Wu, and Loren Dean Williams. 2016. “Yeast rRNA Expansion Segments: Folding and Function.” *Journal of Molecular Biology* 428 (20): 4048–59.
11. Hassouna, N., B. Michot, and J. P. Bachellerie. 1984. “The Complete Nucleotide Sequence of Mouse 28S rRNA Gene. Implications for the Process of Size Increase of the Large Subunit rRNA in Higher Eukaryotes.” *Nucleic Acids Research* 12 (8): 3563–83.
12. Khatter, Heena, Alexander G. Myasnikov, S. Kundhavai Natchiar, and Bruno P. Klaholz. 2015. “Structure of the Human 80S Ribosome.” *Nature* 520 (7549): 640–45.
13. Mauro, Vincent P., and Gerald M. Edelman. 2002. “The Ribosome Filter Hypothesis.” *Proceedings of the National Academy of Sciences of the United States of America* 99 (19): 12031–36.
14. Natchiar, S. Kundhavai, Alexander G. Myasnikov, Hanna Kratzat, Isabelle Hazemann, and Bruno P. Klaholz. 2017. “Visualization of Chemical Modifications in the Human 80S Ribosome Structure.” *Nature* 551 (7681): 472–77.
15. Parker, Michael S., Floyd R. Sallee, Edwards A. Park, and Steven L. Parker. 2015. “Homoiterons and Expansion in Ribosomal RNAs.” *FEBS Open Bio* 5 (October): 864–76.
16. Petrov, Anton S., Burak Gulen, Ashlyn M. Norris, Nicholas A. Kovacs, Chad R. Bernier, Kathryn A. Lanier, George E. Fox, et al. 2015. “History of the Ribosome and the Origin of Translation.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (50): 15396–401.

17. Porter, Douglas F., Weili Miao, Xue Yang, Grant A. Goda, Andrew L. Ji, Laura K. H. Donohue, Maria M. Aleman, Daniel Dominguez, and Paul A. Khavari. 2021. “easyCLIP Analysis of RNA-Protein Interactions Incorporating Absolute Quantification.” *Nature Communications* 12 (1): 1569.
18. Shankar, Vaishnavi, Robert Rauscher, Julia Reuther, Walid H. Gharib, Miriam Koch, and Norbert Polacek. 2020. “rRNA Expansion Segment 27Lb Modulates the Factor Recruitment Capacity of the Yeast Ribosome and Shapes the Proteome.” *Nucleic Acids Research* 48 (6): 3244–56.
19. Shi, Zhen, Kotaro Fujii, Kyle M. Kovary, Naomi R. Genuth, Hannes L. Röst, Mary N. Teruel, and Maria Barna. 2017. “Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-Wide.” *Molecular Cell* 67 (1): 71–83.e7.
20. Simsek, Deniz, Gerald C. Tiu, Ryan A. Flynn, Gun W. Byeon, Kathrin Leppek, Adele F. Xu, Howard Y. Chang, and Maria Barna. 2017. “The Mammalian Ribo-Interactome Reveals Ribosome Functional Diversity and Heterogeneity.” *Cell* 169 (6): 1051–65.e18.
21. Van Nostrand, Eric L., Peter Freese, Gabriel A. Pratt, Xiaofeng Wang, Xintao Wei, Rui Xiao, Steven M. Blue, et al. 2020. “A Large-Scale Binding and Functional Map of Human RNA-Binding Proteins.” *Nature* 583 (7818): 711–19.
22. Veldman, G. M., J. Klootwijk, V. C. de Regt, R. J. Planta, C. Branlant, A. Krol, and J. P. Ebel. 1981. “The Primary and Secondary Structure of Yeast 26S rRNA.” *Nucleic Acids Research* 9 (24): 6935–52.
23. Ware, V. C., B. W. Tague, C. G. Clark, R. L. Gourse, R. C. Brand, and S. A. Gerbi. 1983. “Sequence Analysis of 28S Ribosomal DNA from the Amphibian Xenopus Laevis.” *Nucleic Acids Research* 11 (22): 7795–7817.
24. Yang, Tzu-Hsien, Chung-Yu Wang, Hsiu-Chun Tsai, and Cheng-Tse Liu. 2021. “Human IRES Atlas: An Integrative Platform for Studying IRES-Driven Translational Regulation in Humans.” *Database: The Journal of Biological Databases and Curation* 2021 (May). https://doi.org/10.1093/database/baab025.
25. Yokoyama, Takeshi, and Tsutomu Suzuki. 2008. “Ribosomal RNAs Are Tolerant toward Genetic Insertions: Evolutionary Origin of the Expansion Segments.” *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkn224.
26. Zhao, Jian, Yan Li, Cong Wang, Haotian Zhang, Hao Zhang, Bin Jiang, Xuejiang Guo, and Xiaofeng Song. 2020. “IRESbase: A Comprehensive Database of Experimentally Validated Internal Ribosome Entry Sites.” *Genomics, Proteomics & Bioinformatics* 18 (2): 129–39.